Week 13 Homework

1 预习说明

请同学们从以下链接下载补充教材 “Selected Applications of Convex Optimization”：
https://link.springer.com/book/10.1007/978-3-662-46356-7
其中第二章 Support Vector Machines, 2.1 Basic SVM 是预习内容。**下周小测会考察**。

2 作业

I. **Proximal Gradient Descent** (25 points)
Consider the following LASSO problem:

$$\min_x \frac{1}{2}\|Ax-b\|_2^2 + \|x\|_1.$$

Let $f(x) = \frac{1}{2}\|Ax-b\|_2^2, g(x) = \|x\|_1$ and the objective function $h(x) = f(x) + g(x)$. Since f has Lipschitz gradient, we have

$$\begin{aligned}
&x^* = \arg\min_x f(x) + g(x) \\
\Leftrightarrow &0 \in (\nabla f + \partial g)x^* \\
\Leftrightarrow &0 \in (I - \alpha\nabla f)x^* - (I + \alpha\partial g)x^* \\
\Leftrightarrow &(I - \alpha\nabla f)x^* \in (I + \alpha\partial g)x^* \\
\Leftrightarrow &(I + \alpha\partial g)^{-1}(I - \alpha\nabla f)x^* = x^* \\
\Leftrightarrow &x^* = \mathrm{Prox}_{\alpha g} \circ (I - \alpha\nabla f)x^*,
\end{aligned}$$

where $\circ$ is the composition of operators.

(1) [4 points] Fix data A, b, prove that $\nabla f(x)$ is Lipschitz and show the smallest positive constant scalar M such that $\|\nabla f(x) - \nabla f(y)\|_2 \le M\|x-y\|$, $\forall x, y$. (show M as an expression of A, b)

(2) [5 points] Write out the exact form of iteration

$$x_{k+1} = \text{Prox}_{\alpha g} \circ (I - \alpha \nabla f) x_k.$$

Choose $\alpha = 1/M$, achieve the iteration by coding.

(3) [12 points] Given two groups of A, b in files `A1.csv, b1.csv` (path `./problem1data/`) where $A1$ is full-rank and `A2.csv, b2.csv` where $A2$ is not full-rank, implement the iteration to solve the LASSO problem. Start with $x_0 = \mathbf{0}$. Stop when $|h(x_{k+1}) - h(x_k)| < 10^{-5}$. Plot the corresponding figure of $\log(k)$ vs $\log(\|x_k - x^*\|_2)$ and $\log(k)$ vs $\log(h(x_k))$ and $\log(k)$ vs $\log(h(x_{k-1}) - h(x_k))$. (6 figures in total)

(4) [4 points] Discuss the convergence rate difference between two groups of data and explain by referring to convergence theories introduced in lecture.

II. **(选做题) Douglas-Rachford Splitting (DRS) and Alternating Direction Method of Multipliers (ADMM)**

Consider the following optimization problem

$$\min_x \|x\|_1, \text{ s.t. } Ax = b$$

- The data A, b is given in files `A.csv, b.csv`, in path `./problem2data/`
- To plot the curve, all the algorithms stop when $\|x_{k-1}\|_1 - \|x_k\|_1 < 10^{-5}$.

1.(+7 points) Reformulate the problem as

$$\min_x \|x\|_1 + \mathbb{I}_{\{Ax=b\}}(x).$$

Define $f(x) = \|x\|_1, g(x) = \mathbb{I}_{\{Ax=b\}}(x)$. The iteration given by DRS is

$$\begin{aligned} x_{k+1} &= \text{Prox}_{\alpha g}(z_k), \\ y_{k+1} &= \text{Prox}_{\alpha f}(2x_{k+1} - z_k), \\ z_{k+1} &= z_k + y_{k+1} - x_{k+1}. \end{aligned}$$

Write the exact form of this iteration and achieve the iteration by coding. Choose appropriate α. Given A, b in files, implement the iteration to solve the optimization problem. Plot the corresponding figure of $\log(k)$ vs $\log(\|x_k - x^*\|_2)$ and $\log(k)$ vs $\log(\|x_k\|_1)$.

2.(+8 points) Reformulate the problem as

$$\min_x \|x\|_1, \text{ s.t. } x = y, Ay = b.$$

Define the augmented Lagrangian as

$$L_\alpha(x,y,u,v) = \|x\|_1 + u^T(x-y) + v^T(Ay-b) + \frac{\alpha}{2}\|x-y\|_2^2 + \frac{\alpha}{2}\|Ay-b\|_2^2$$

The ADMM iteration is

$$\begin{aligned}
x_{k+1} &\in \arg\min_x \ L_\alpha(x, y_k, u_k, v_k),\\
y_{k+1} &\in \arg\min_y \ L_\alpha(x_{k+1}, y, u_k, v_k),\\
u_{k+1} &= u_k + \alpha(x_{k+1} - y_{k+1}),\\
v_{k+1} &= v_k + \alpha(Ay_{k+1} - b).
\end{aligned}$$

Derive the exact form of this iteration and achieve the iteration by coding. Choose appropriate α. Given A, b in files, implement the iteration to solve the optimization problem. Plot the corresponding figure of $\log(k)$ vs $\log(\|x_k - x^*\|_2)$ and $\log(k)$ vs $\log(\|x_k\|_1)$.

3.(+10 points) Reformulate the problem as

$$\min_x \|x\|_1, \ \text{s.t.} \ Ax = b, y = Px.$$

Derive the iteration of ADMM. Select P such that the x update can be completed via soft-thresholding. Hints can be found in Lecture Notes.

Achieve the iteration by coding. Choose appropriate update step size. Given A, b in files, implement the iteration to solve the optimization problem. Plot the corresponding figure of $\log(k)$ vs $\log(\|x_k - x^*\|_2)$ and $\log(k)$ vs $\log(\|x_k\|_1)$.

3 作业说明

- **注意：本次作业必做题 25 分（与之前周一致），选做题 25 分，最多可得 50 分。**
- **所有题目需要理论推导和过程，只画图即使对了也只能得部分分数。**报告提交电子版，和代码一起打包提交至网络学堂。提交作业时文件夹中应包含数据文件，保证程序可以直接在文件夹中运行。
- 本次作业截止时间下周六晚上 23：59，**不接受补交作业。**